A unifying formulation of the Fokker-Planck-Kolmogorov equation for general stochastic hybrid systems (extended version)

Julien Bect

A unifying formulation of the Fokker-Planck-Kolmogorov equation for general stochastic hybrid systems[☆,☆☆]

Julien Bect

Department of Signal Processing and Electronic Systems, Supelec, Gif-sur-Yvette, France.

Abstract

A general formulation of the Fokker-Planck-Kolmogorov (FPK) equation for stochastic hybrid systems is presented, within the framework of Generalized Stochastic Hybrid Systems (GSHSs). The FPK equation describes the time evolution of the probability law of the hybrid state. Our derivation is based on the concept of mean jump intensity, which is related to both the usual stochastic intensity (in the case of spontaneous jumps) and the notion of probability current (in the case of forced jumps). This work unifies all previously known instances of the FPK equation for stochastic hybrid systems, and provides GSHS practitioners with a tool to derive the correct evolution equation for the probability law of the state in any given example.

Key words: Stochastic hybrid systems, Stochastic system with jumps, Markov processes, Fokker-Planck equation

1. Introduction

Among all continuous-time stochastic models of (nonlinear) dynamical systems, those with the Markov property are especially appealing because of their numerous nice properties. In particular, they come equipped with a pair of operator semigroups, the so-called backward and forward semigroups, which are the analytical keys to many practical problems involving Markov processes. When the system is determined by a stochastic differential equation, these semigroups are generated by Partial Differential Equations (PDE) — respectively the backward and forward Kolmogorov equations. The forward Kolmogorov PDE, also known as the Fokker-Planck equation, rules the time evolution $t \mapsto \mu_t$, where μ_t is the probability distribution of the state X_t of the system at time t. This paper deals with the generalization of this Fokker-Planck-Kolmogorov (FPK) equation to the framework of General Stochastic Hybrid Systems (GSHSs) recently proposed by Bujorianu and Lygeros [7, 8].

The GSHS framework encompasses nearly all continuous-time Markov models arising in practical applications, including piecewise deterministic Markov processes [11, 12], switching diffusions [18, 19] and the stochastic hybrid systems of Hu et al. [21]. The reader is referred to [9, 31] for a detailed overview of these classes of models with a view towards applications in Air Traffic Management (ATM). Two kinds of jumps are allowed in a GSHS: spontaneous jumps, defined by a state-dependent stochastic intensity $\lambda(X_t)$, and forced jumps triggered by a so-called guard set G. Generalized FPK equations have been given in the literature, in the case of spontaneous jumps, for several classes of models; see Gardiner [17], Kontorovich and Lyandres [23], Krystul et al. [24] and Hespanha [20] for instance. The case of forced jumps is harder to analyze, at the FPK level, because no stochastic intensity exists for these jumps. Until recently, the only results available in the literature were dealing with one-dimensional models; see Feller [15, 16] and Malhamé and Chong [28]. These results have been extended to a class of multi-dimensional models by Bect et al. [3].

The main contribution of this paper is a general formulation of the FPK equation for GSHSs. It is based on the concept of *mean jump intensity*, which conveniently substitutes for the stochastic intensity when the latter does not exist. This equation unifies all previously known instances of the FPK equation for stochastic hybrid systems, and

[☆] A shorter version of this paper was presented at the 17th IFAC World Congress (IFAC'08) in Seoul, Korea [2].

[☆☆] The results presented in this paper come from the PhD thesis of the author [1], under the supervision of Pr. Gilles Fleury and Dr. Hana Baili.

Email addresses: `julien.bect@supelec.fr` (Julien Bect)

Paper published by Elsevier in Nonlinear Analysis: Hybrid Systems, 4(2):357–370, may 2010. This is an author-generated post-print version.

provides GSHS practitioners with a tool to derive the correct evolution equation for the probability law of the state in any given example. The results presented in this paper are extracted from the PhD thesis of the author [1].

The paper is organized as follows. Section 2 introduces our notations for the GSHS formalism, together with various assumptions that will be needed in what follows. Section 3 defines the crucial concept of mean jump intensity, which is used in Section 4 to derive our unified measure-valued formulation of the generalized FPK equation for GSHSs. Section 5 shows that the measure-valued equation of Section 4 yields an evolution equation with associated boundary conditions in the case where a piecewise smooth exists. Section 6 provides several examples, showing that the generalized FPK equation allows to recover all known instances of the FPK equation for stochastic hybrid systems. Finally, Section 7 concludes the paper and provides directions for future work.

Remark 1. The stochastic processes that we call GSHSs, following the papers by Bujorianu and Lygeros [7, 8], are also called GSHPs in [9] — where GSHP stands for "General Stochastic Hybrid Process". Note also that the terms GSHS / GSHP are used by [14] to make a clear distinction between the formal data defining the process and the process itself. We shall stick to the "GSHS" terminology in this paper.

2. General Stochastic Hybrid Systems

From the probabilistic point of view, the object of interest in the GSHS formalism is a continuous-time strong Markov process $X = (X_t)_{t\geq 0}$, with values in a metric space E^0. It is defined on a filtered space $(\Omega, \mathcal{A}, \mathcal{F})$, equipped with a system $\left\{\boldsymbol{P}_x;\, x \in E^0\right\}$ of probability measures on $(\Omega, \mathcal{A})$, with the property that X starts from x under $\boldsymbol{P}_x$ for all $x \in E^0$. As usual in the theory of Markov processes, $\boldsymbol{E}_x$ denotes the expectation operator corresponding to $\boldsymbol{P}_x$. The reader is referred to [6, 13, 34] for background information on continuous-time Markov processes.

It is assumed that, for each $\omega \in \Omega$, the samplepath $t \mapsto X_t(\omega)$ is right-continuous, has left limits $X_t^-(\omega)$ in the completion E of E^0, and has a finite number of jumps, denoted by $N_t(\omega)$, on the interval $(0;t]$ for all $t \geq 0$. The last condition can be seen as a "pathwise non-Zenoness" requirement. We will denote by τ_k the k^{th} jump time, with $\tau_k = +\infty$ if there is less than k jumps.

2.1. *The hybrid state space*

The (completed) state space of the model is assumed to have a hybrid structure: $E = \cup_{q\in\mathcal{Q}} \{q\} \times E_q$, where $\mathcal{Q}$ is a finite or countable set, and each E_q is either the closure of some connected open subset $D_q \subset \mathbb{R}^{n_q}$ ($n_q \geq 1$) or a singleton (in which case we set $n_q = 0$). The state at time t can therefore be written as a pair $X_t = (Q_t, Z_t)$, where $Q_t \in \mathcal{Q}$ and $Z_t \in E_{Q_t}$. We denote by $\mathcal{Q}^{\mathrm{d}} = \left\{q \in \mathcal{Q} \,\middle|\, n_q = 0\right\}$ the set of all "purely discrete" modes, and by $E^{\mathrm{d}} = \cup_{q\in\mathcal{Q}^{\mathrm{d}}} \{q\} \times E_q$ the corresponding subset of E. The usual definitions for smooth maps and vector fields extend without difficulty to such a hybrid structure (see Appendix A for details).

The state space E is regarded as the disjoint sum of the sets E_q, $q \in \mathcal{Q}$, and endowed with the disjoint union topology[1]. We denote by $\mathcal{E}$ the Borel σ-algebra, and by $\mathcal{E}_{\mathrm{c}}$ the subset of all relatively compact $\Gamma \in \mathcal{E}$. Moreover, we define a "volume measure" on E by the relation

$$\mathfrak{m}(\Gamma) = \sum_{q\notin\mathcal{Q}^{\mathrm{d}}} \mathfrak{m}_q(\Gamma \cap E_q) + \sum_{x\in E^{\mathrm{d}}} \delta_x(\Gamma)\,, \quad \Gamma \in \mathcal{E}, \tag{1}$$

where $\mathfrak{m}_q$ is the n_q-dimensional Lebesgue measure on E_q and δ_x the Dirac mass at x. (Note that $E_q \subset \mathbb{R}^{n_q}$ has been tacitly identified with $\{q\} \times E_q \subset E$.)

Let ∂E_q be the boundary of E_q in $\mathbb{R}^{n_q}$, with the convention that $\partial E_q = \varnothing$ when $n_q = 0$. We define the boundary ∂E of the state space by the relation $\partial E = \cup_{q\in\mathcal{Q}} \{q\} \times \partial E_q$, and the *guard set* by $G = E \setminus E^0$. It is *not* required that $G = \partial E$.

Notations. Let $\mu : \mathcal{E} \to \mathbb{R}$ be a (signed) measure, $K : E \times \mathcal{E} \mapsto \mathbb{R}$ a kernel and $\varphi : E \to \mathbb{R}$ a measurable function. The following notations will be used throughout the paper, assuming the integrals exist: $(\mu K)(\mathrm{d}y) = \int \mu(\mathrm{d}x)\, K(x, \mathrm{d}y)$, $(K\varphi)(x) = \int K(x, \mathrm{d}y)\, \varphi(y)$ and $\mu\varphi = \int \mu(\mathrm{d}x)\, \varphi(x)$.

[1] which is (here) locally compact, separable and completely metrizable

2.2. Stochastic differential equation with jumps

The process X is assumed to be driven by an Itô stochastic differential equation between its jumps: there exist $r+1$ smooth vector fields $\mathbf{f}_l$ and a r-dimensional Wiener process B such that, in mode $q \in \mathcal{Q} \setminus \mathcal{Q}^{\mathrm{d}}$,

$$\mathrm{d}Z_t \;=\; \mathbf{f}_0(q, Z_t)\,\mathrm{d}t + \sum_{l=1}^{r} \mathbf{f}_l(q, Z_t)\,\mathrm{d}B_t^l\,. \tag{2}$$

In other words, for all $\varphi \in C^2(E)$, X satisfies the following generalized Itô formula

$$\varphi(X_t) - \varphi(X_0) = \int_0^t (L\varphi)(X_s)\,\mathrm{d}s + \sum_{l=1}^{r} \int_0^t (\mathbf{f}_l\varphi)(X_s)\,\mathrm{d}B_s^l + \sum_{0<\tau_k\le t} \bigl(\varphi(X_{\tau_k}) - \varphi(X_{\tau_k}^-)\bigr)\,, \tag{3}$$

where L is the differential generator associated with (2), i.e.

$$L = \sum_i \mathbf{f}_0^i \frac{\partial}{\partial z^i} + \tfrac{1}{2} \sum_{i,j} \left(\sum_{l=1}^{r} \mathbf{f}_l^i \mathbf{f}_l^j \right) \frac{\partial^2}{\partial z^i \partial z^j}\,. \tag{4}$$

We make the following smoothness assumptions:

Assumption 2. *The drift $\mathbf{f}_0$ is of class C^1, and the other vector fields $\mathbf{f}_l$, $1 \le l \le r$, are of class C^2.*

Remark 3. It would be possible to slightly generalize the model by considering a mode-dependent number r_q of Wiener processes. All the results of the paper would still hold with the same proofs. We choose to use the same number r of Wiener processes in each mode for the sake of notational simplicity. Note that this is consistent with the most recent definitions of GSHS [8, 14], but not with the one in [7] (which uses a mode-dependent number of noise processes).

2.3. Two different kinds of jumps

We assume that there exists a Markov kernel K from E to E^0 and a measurable locally bounded function $\lambda : E^0 \to \mathbb{R}_+$, such that the following *Lévy system identity* holds for all $x \in E^0$, $t \ge 0$, and for all measurable $\varphi : E \times E^0 \to \mathbb{R}_+$:

$$\boldsymbol{E}_x \left\{ \textstyle\sum_{0<\tau_k\le t} \varphi(X_{\tau_k}^-, X_{\tau_k}) \right\} = \boldsymbol{E}_x \left\{ \int_0^t (K\varphi)(X_s^-)\,\mathrm{d}H_s \right\} \tag{5}$$

where $(K\varphi)(y) = \int_{E^0} K(y, \mathrm{d}y')\,\varphi(y, y')$ and H is the predictable increasing process defined by

$$H_t = \int_0^t \lambda(X_s)\,\mathrm{d}s + \sum_{\tau_k \le t} \mathbb{1}_{X_{\tau_k}^- \in G}\,. \tag{6}$$

The first part corresponds to *spontaneous* jumps, triggered "randomly in time" with a stochastic intensity $\lambda(X_t)$, while the other part corresponds to *forced* jumps, triggered when X hits the guard set G.

Remark 4. The terms "spontaneous" and "forced" seem to have been coined by Bujorianu et al. [9]. They are closely related to the probabilistic notions of predictability and total inaccessibility for stopping times [see, e.g., 32, chapter VI, §§12–18], but we shall not discuss this point further in this paper.

Remark 5. The pair (K, H) is a *Lévy system* for the process X in the sense of Walsh and Weil [35, definition 6.1]. Most authors require that H be continuous in the definition of a Lévy system, thereby disallowing predictable jumps.

3. Mean jump intensity

From now on, we assume that some initial probability law μ_0 has been chosen, with $\mu_0(G) = 0$ since the process cannot start from G. All expectations will be taken, without further mention, with respect to the probability $\boldsymbol{P}_{\mu_0} = \int \mu_0(\mathrm{d}x)\boldsymbol{P}_x$.

It is assumed that $\boldsymbol{E}(N_t) < +\infty$. This is a usual requirement for stochastic hybrid processes[2], which is clearly stronger than piecewise-continuity of the samplepaths. Its being satisfied depends not only on the dynamics of the system but also on the initial probability law μ_0.

[2]See, e.g., Davis [11] or Bujorianu and Lygeros [7].

3.1. *Definition and connection with the usual stochastic intensity*

In order to introduce the main concept of this section, let us define a positive measure R on $E \times (0; +\infty)$ by

$$R\,(A) = \boldsymbol{E}\Big\{ \sum\nolimits_{k \geq 1} \mathbb{1}_A \left(X_{\tau_k}^{-}, \tau_k \right) \Big\}\,. \tag{7}$$

For any $\Gamma \in \mathcal{E}$, the quantity $R\,(\Gamma \times (0; t])$ is the expected number of jumps starting from Γ during the time interval $(0; t]$. The measure R is in general unbounded, but its restriction to $E \times (0; t]$ is bounded for all $t \geq 0$ because $\boldsymbol{E}(N_t) < +\infty$.

Definition 6. Suppose that there exists a mapping $r : t \mapsto r_t$, from $[0; +\infty)$ to the set of all positive bounded measures on E, such that, for all $\Gamma \in \mathcal{E}$,

a) $t \mapsto r_t(\Gamma)$ is measurable,

b) for all $t \geq 0$, $R\,(\Gamma \times (0; t]) = \int_0^t r_s(\Gamma)\, \mathrm{d}s$.

Then r is called the *mean jump intensity* of the process X (started with the initial law μ_0).

Let us split R into the sum of two measures R^0 and R^G, corresponding respectively to the spontaneous and forced jumps of the process. Then, using the Lévy system identity, it is easy to see that a mean jump intensity r^0 always exist for the spontaneous part R^0: it is given by

$$r_t^0(\Gamma) = \boldsymbol{E}\big(\lambda(X_t)\, \mathbb{1}_{X_t \in \Gamma}\big) = \int_\Gamma \lambda(x)\, \mu_t(\mathrm{d}x)\,. \tag{8}$$

In other words: for spontaneous jumps, a mean jump intensity always exists, and it is the expectation of the stochastic jump intensity $\lambda(X_t)$ on the event $\{X_t \in \Gamma\}$.

Forced jumps are more problematic. The Lévy system identity is powerless here, since no stochastic intensity exists (because forced jumps are predictable). All hope is not lost, though: a simple example will be presented in the next subsection, proving that a mean jump intensity can exist anyway. This is fortunate, since the existence of a mean jump intensity will be an essential ingredient for our unified formulation of the generalized FPK equation. See subsection 6.2 for further details on that issue.

3.2. *Where μ_0 comes into play: an illustrative example*

Consider the following hybrid dynamics on $E = [0; 1]$: the state X_t moves to the right at constant speed $v > 0$ as long as it is in $E^0 = [0; 1)$, and jumps instantaneously to 0 as soon as it hits the guard $G = \{1\}$ (i.e., the reset kernel is such that $K(1, \cdot\,) = \delta_0$).

If we take $\mu_0 = \delta_0$ for the initial law, then the process jumps from 1 to 0 each time t is a multiple of $1/v$, i.e. $\tau_k = k/v$ and $X_{\tau_k}^{-} = 1$ almost surely. There is therefore no mean jump intensity in this case, since $R = \sum_{k \geq 1} \delta_{(1,\, k/v)}$.

Now take μ_0 to be the uniform probability on $[0; 1]$ (which is, incidentally, the only stationary probability law of the process). Then

$$\begin{aligned} R\big(\Gamma \times (0; t]\big) \;&=\; \delta_1(\Gamma) \int_0^1 \operatorname*{argmax}_{k \geq 1} \left\{ \frac{k - x}{v} \leq t \right\} \mathrm{d}x && (9)\\ &=\; \delta_1(\Gamma) \int_0^1 \lceil vt + x \rceil \;\mathrm{d}x && (10)\\ &=\; vt\, \delta_1(\Gamma)\,, && (11) \end{aligned}$$

where $\lceil vt + x \rceil$ is the smaller integer greater or equal to $vt + x$. Therefore the mean jump intensity exists in this case, and is equal to $v\, \delta_1$ (it is of course time-independent, since μ_0 is stationary). In particular, the global mean jump intensity is $r_t(E) = v$.

4. Generalized FPK equation

4.1. A weak form of the FPK equation

Taking expectations in (3), the following *generalized Dynkin formula* is obtained: for all compactly supported $\varphi \in C^2(E)$ and all $t \geq 0$,

$$\boldsymbol{E}\left\{\varphi(X_t) - \varphi(X_0)\right\} \;=\; \boldsymbol{E}\left\{\int_0^t (L\varphi)(X_s)\,\mathrm{d}s\right\} + \boldsymbol{E}\left\{\sum_{0<\tau_k\leq t} \varphi(X_{\tau_k}) - \varphi(X_{\tau_k}^-)\right\}. \tag{12}$$

Let us assume the existence of a mean jump intensity r_t at all times. Then (12) can be rewritten as

$$(\mu_t - \mu_0)\,\varphi \;=\; \int_0^t \mu_s(L\varphi)\,\mathrm{d}s + \int_0^t r_s(K-I)\varphi\,\mathrm{d}s\,, \tag{13}$$

where μ_t is the law of X_t and I is the "identity kernel" on E, i.e. the kernel defined by $I(y, \mathrm{d}y') = \delta_y(\mathrm{d}y')$. Formally differentiating (13) yields

$$\mu_t' \;=\; L^*\mu_t + r_t(K-I)\,, \tag{14}$$

where $t \mapsto \mu_t'$ is the time derivative of $t \mapsto \mu_t$ (in a sense to be specified later), and L^* the "distributional adjoint" of L, defined over the set $\mathcal{M}_\mathrm{c}(E)$ of all signed Radon measures[3] on E by

$$(L^*\nu)\,(\varphi) \;=\; \nu\,(L\varphi) \;=\; \int_E (L\varphi)(x)\,\nu(\mathrm{d}x)\,, \qquad \forall\nu \in \mathcal{M}_\mathrm{c}(E),\ \forall\varphi \in C_\mathrm{c}^2(E)\,. \tag{15}$$

As a consequence of Assumption 2, the result $L^*\nu$ of applying L^* to a Radon measure ν is, in general, a second-order distribution. It is important to note that, because the state space E has a boundary ∂E, the operator L^* is not a simple second-order partial differential operator – it also includes "boundary terms".

Equation (14) begins like the usual Fokker-Planck equation for diffusion processes ($\mu_t' = L^*\mu_t$) and ends with an additional term that accounts for the jumps of the process.

Definition 7. We will say that $t \mapsto \mu_t$ is a solution in the weak sense of the *generalized FPK equation* for the GSHS if

a) there exists a mean jump intensity $t \mapsto r_t$,

b) there exists a mapping $t \mapsto \mu_t'$, from $[0; +\infty)$ to $\mathcal{M}_\mathrm{c}(E)$, such that $t \mapsto \mu_t(\Gamma)$ is absolutely continuous with a.e.-derivative $t \mapsto \mu_t'(\Gamma)$, for all $\Gamma \in \mathcal{E}_\mathrm{c}$,

c) $L^*\mu_t$ is a Radon measure for all $t \geq 0$,

d) equation (14) holds as an equality between Radon measures, i.e. $\mu_t'(\Gamma) = (L^*\mu_t)(\Gamma) + r_t(K-I)(\Gamma)$ for all $t \geq 0$ and all $\Gamma \in \mathcal{E}_\mathrm{c}$.

Such a weak form of the FPK equation is the price to pay for a unified treatment of both kind of jumps. Conditions 7.a and 7.b can be seen as smoothness requirements with respect to the time variable, and 7.c with respect to the space variables.

[3] In this paper, a "Radon measure" will always be a *signed* Radon measure, in other words a distribution of order zero; see Rudin [33] for the basic definitions and properties of distributions. Any signed Radon measure ν can be written as the difference $\nu = \nu^+ - \nu^-$ of two *positive* Radon measures (i.e. locally finite measures); see, e.g., Cohn [10, chapter 7] for more information on the connection between the functional analytic and the measure-theoretic point of view.

4.2. "Physical" interpretation

The usual FPK equation admits a well-known physical interpretation as a conservation equation for the "probability mass" [see, e.g., 17]. Indeed, assuming the existence of a smooth pdf $p \in C^{2,1}(E \times \mathbb{R}_+)$, the equation $\mu_t' = L^*\mu_t$ can be rewritten as a conservation equation $\partial p_t / \partial t + \mathrm{div}(\mathbf{j}_t) = 0$, with the *probability current* $\mathbf{j}_t$ defined by

$$\mathbf{j}_t^i = \mathbf{f}_0^i \, p_t - \frac{1}{2} \sum_j \frac{\partial (a^{ij} p_t)}{\partial z^j}\,, \quad a^{ij} = \sum_{l=1}^{r} \mathbf{f}_l^i \mathbf{f}_l^j\,. \tag{16}$$

The additional "jump term", in the generalized FPK equation, admits a nice physical interpretation as well. To see this, let us rewrite it as the difference of two bounded positive measure: $r_t(K - I) = r_t^{\mathrm{src}} - r_t$, where $r_t^{\mathrm{src}} = r_t K$. Therefore r_t and r_t^{src} behave respectively as a *sink* and a *source* in the generalized FPK equation: for each $\Gamma \in \mathcal{E}$, $r_t(\Gamma)\,\mathrm{d}t$ is the probability mass leaving the set Γ during $\mathrm{d}t$, because of the jumps of the process, while $r_t^{\mathrm{src}}(\Gamma)\,\mathrm{d}t$ is the probability mass entering Γ.

These two measures are in fact connected by the reset kernel $K(x, \mathrm{d}y)$. In particular, the relation $r_t(E) = r_t^{\mathrm{src}}(E)$ holds at all times $t \geq 0$, ensuring that the total probability mass is conserved. Moreover, introducing the measures $W_t(\mathrm{d}x, \mathrm{d}y) = r_t(\mathrm{d}x) K(x, \mathrm{d}y)$, we have $r_t = \int W(\cdot, \mathrm{d}x)$, $r_t^{\mathrm{src}} = \int W(\mathrm{d}x, \cdot)$ and the generalized FPK equation can be rewritten more symmetrically as

$$\mu_t' \;=\; L^* \mu_t + \int \left(W_t(\mathrm{d}x, \cdot) - W_t(\cdot, \mathrm{d}x) \right)\,. \tag{17}$$

It appears clearly, under this form, as a generalization of the *differential Chapman-Kolmogorov formula* of Gardiner [17, equation 3.4.22] — which only allows spontaneous jumps.

4.3. Sufficient conditions for the existence of a weak solution

The main result of this paper shows that the various requirements of definition 7 are not independent. We denote by $|\nu|$ the total variation measure of a Radon measure ν, which is finite on $\mathcal{E}_{\mathrm{c}}$. We shall say that a function $t \mapsto \nu_t$ from $[0; \infty)$ to $\mathcal{M}_{\mathrm{c}}(E)$ is right-continuous (resp. locally integrable) if $t \mapsto \nu_t \varphi$ is right-continuous (resp. locally integrable) for all bounded measurable $\varphi : E \to \mathbb{R}$.

Theorem 8. *Consider the following assumptions:*

a) there exists a mean jump intensity r, such that $t \mapsto r_t$ is right-continuous,

b) $t \mapsto \mu_t$ is differentiable in the sense of 7.b, $t \mapsto \mu_t'$ is right-continuous and $t \mapsto |\mu_t'|$ locally integrable,

c) $L^*\mu_t$ is a Radon measure for all $t \geq 0$, $t \mapsto L^*\mu_t$ is right-continuous and $t \mapsto |L^*\mu_t|$ is locally integrable.

If any two of these assumptions hold, then the third holds as well and $t \mapsto \mu_t$ is a solution in the weak sense of the generalized FPK equation.

The proof of this theorem is given in Appendix B. We will not try to give general conditions under which assumptions 8.a–8.c are satisfied, since such conditions would inevitably be, in the general setting of this paper, very complicated (involving the initial law μ_0, the vector fields $\mathbf{g}$ of the stochastic differential equation, the geometry of the state space E and the reset kernel K).

5. The case when a piecewise smooth pdf exists

Equation (14) is an evolution equation for the measure-valued function $t \mapsto \mu_t$. In many situations of practical interest, the measures μ_t admit a pdf p_t, with respect to the volume measure $\mathfrak{m}$ on E. In this section we show that, if the function $p : (x, t) \mapsto p_t(x)$ is – at least piecewise – smooth, then equation (14) simultaneously gives birth to an evolution equation for $t \mapsto p_t$ and to static relations that hold for all $t \geq 0$ (the so-called "boundary conditions", although the name is not entirely appropriate here).

5.1. Assumptions about the guard and the boundary

Turning equation (14) into an evolution equation for the pdf ultimately boils down to playing with “integration by parts” formulas, for judiciously chosen test functions. To do so, we shall need additional assumptions concerning the topological regularity of the guard set and the smoothness of the boundary.

Assumption 9. *The guard set G is a regular closed subset of ∂E (i.e., G is a closed set and it is equal to the closure of its interior in ∂E).*

Assumption 10. *For each $q \in \mathcal{Q}$ such that $n_q \geq 2$, the domain E_q is C^2-manifold with corners.*

See Lee [26, chapter 14] for basic definitions and results concerning manifolds with corners. Assumption 10 is sufficient for the divergence theorem to hold (see Appendix A for a precise statement). The divergence theorem is a multi-dimensional generalization of the “integration by parts” formula, and will be the key tool to compute $L^*\nu$ for Radon measures with a smooth density.

We denote by $\mathfrak{s}_q$ the surface measure on ∂E_q, and define the surface measure $\mathfrak{s}$ on ∂E by

$$\mathfrak{s} \;=\; \sum_{\substack{q\in\mathcal{Q}\\ n_q\geq 2}} \mathfrak{s}_q \;+\; \sum_{\substack{q\in\mathcal{Q}\\ n_q=1}} \sum_{x\in\partial E_q} \delta_x \,. \tag{18}$$

We further denote by $\mathbf{n}$ the outward-pointing unit normal vector on ∂E, which is well-defined $\mathfrak{s}$-almost everywhere on ∂E. Since the process X is allowed to start on $\partial E \setminus G$, which is a subset of E^0 (see Section 2), the vector fields have to satisfy the following conditions (on the smooth part of $\partial E \setminus G$, hence $\mathfrak{s}$-almost everywhere):

$$\langle \mathbf{f}_0, \mathbf{n}\rangle \leq 0\,, \quad \text{and} \quad \langle \mathbf{f}_l, \mathbf{n}\rangle = 0\,,\; 1 \leq l \leq r\,. \tag{19}$$

Otherwise, for any $(q,x) \in \partial E \setminus G$, the solution of equation (2) would leave the domain “instantaneously” (i.e. almost surely in any time neighborhood of 0).

5.2. Connecting the mean intensity of forced jumps with the probability current (local result)

Let G^0 denote the subset of the guard set G where at least one of the “noise” vector fields is not tangent to the boundary, i.e. $G^0 = \{x \in G,\ \exists l \in \{1,\ldots,r\},\ \langle \mathbf{f}_l, \mathbf{n}\rangle \neq 0\}$. The following results relates the mean intensity of forced jumps with the probability current $\mathbf{j}_t$ defined by equation (16).

Proposition 11. *Assume that the measures μ_t admit a pdf $p_t = p(\,\cdot\,,t)$ for all $t \geq 0$ on some open subset $U \subset E$, with $p \in C^{2,1}(U \times \mathbb{R}_+)$. Define the outward probability current $j_t^{\mathrm{out}} = \langle \mathbf{j}_t, \mathbf{n}\rangle$ on $U \cap G$. Then, for all $t \geq 0$,*

a) $j_t^{\mathrm{out}} \geq 0$ and $r_t^G(\Gamma) = \int_\Gamma j_t^{\mathrm{out}}\, \mathrm{d}\mathfrak{s}$ is the mean intensity of forced jumps on $U \cap G$,

b) the pdf p_t vanishes on $U \cap G^0$.

See Appendix C for the proof. This proposition provides two important conclusions concerning forced jumps. The first one is that, when a smooth pdf exists in a neighborhood of the guard set, the mean intensity of forced jumps (which appears in the FPK equation) is equal to the outward flow of the probability current. This is consistent with the physical interpretation of the probability current: $\mathbf{j}_t\, \mathrm{d}\mathfrak{s}\, \mathrm{d}t$ is the probability mass (“number of particles”) escaping from the domain through $\mathrm{d}\mathfrak{s}$ during $\mathrm{d}t$.

The second conclusion is that the familiar “absorbing boundary” condition $p_t = 0$ holds on the guard set as soon as one of the “noise” vector fields is active in the normal direction. Note that the pdf does not vanish on the boundary in the example of subsection 3.2, which is a piecewise deterministic process with forced jumps. A “physical” explanation of absorbing boundaries, in the spirit of subsection 4.2, can be found in [28] and also, more recently, in [27].

5.3. Evolution equation for the pdf and “boundary” conditions (global result)

The local result of subsection 5.2 will now be used to obtain a general formulation of the FPK equation (14) in terms of a probability density function, when one exists and is smooth enough. Let $H \subset E^0 \setminus E^{\mathrm{d}}$ be a closed set of $\mathfrak{m}$-measure zero – typically, H will be a closed hypersurface in applications. Note that $U = E \setminus H$ is an open neighborhood of the boundary ∂E. Assume now that the following holds:

Assumption 12. *a)* μ_t *admits a pdf* p_t *with respect to* $\mathfrak{m}$*, on the whole state space, for all* $t \geq 0$*,*

b) $p \in C^{2,1}(U \times \mathbb{R}_+)$*, with* $\frac{\partial p}{\partial t}$ *and* Fp *locally integrable on* $E \times \mathbb{R}_+$*.*

Then, it follows from the proof of Proposition 11 (see Appendix C, equation (48)) that

$$(L^* \mu_t)(\Gamma) \;=\; \int_\Gamma F p_t \,\mathrm{d}\mathfrak{m} \;+\; \int_{\partial E \cap \Gamma} j^{\mathrm{out}} \,\mathrm{d}\mathfrak{s}\,, \qquad \forall \Gamma \in \mathcal{E}_{\mathrm{c}} \text{ such that } \Gamma \subset U\,, \tag{20}$$

where F is the formal adjoint of L, i.e., the differential operator defined by

$$F:\; q \;\mapsto\; -\sum_i \frac{\partial\left(\mathbf{f}_0^i\, q\right)}{\partial z^i} \;+\; \frac{1}{2}\sum_{i,j} \frac{\partial^2\left(a^{ij}\, q\right)}{\partial z^i \partial z^j}\,. \tag{21}$$

The (possible) lack of differentiability of p_t on H therefore translates into the fact that the Radon measures β_t,

$$\beta_t(\Gamma) \;=\; \int_\Gamma F p_t \,\mathrm{d}\mathfrak{m} \;+\; \int_{\partial E \cap \Gamma} j^{\mathrm{out}} \,\mathrm{d}\mathfrak{s} \;-\; (L^* \mu_t)(\Gamma)\,. \tag{22}$$

do not vanish in general. This, in turn, is closely related to the existence of a non-vanishing $\mathfrak{m}$-singular part (see Appendix D for a definition) in the source term $r_t^{\mathrm{src}} = r_t K$, as stated by the following result.

Theorem 13. *Let Assumption 12 hold. Then the conditions 8.a–8.c of Theorem 8 are satisfied, and the following evolution equation holds on* $E^0 \setminus H$*, for all* $t \geq 0$*:*

$$\frac{\partial p}{\partial t} \;=\; F p_t + \frac{\mathrm{d}(r_t K)}{\mathrm{d}\mathfrak{m}} - \lambda\, p_t\,. \tag{23}$$

Moreover, according to Proposition 11,

a) $r_t^G(\Gamma) \;=\; \int_{\Gamma \cap G} j_t^{\mathrm{out}} \,\mathrm{d}\mathfrak{s}$ *is the mean intensity of forced jumps,*

b) and the absorbing boundary condition, $p_t = 0$*, holds on* G^0

Finally, the $\mathfrak{m}$*-singular part* $(r_t K)^\perp$ *of* $r_t K$ *is supported by the set* $H \cup (\partial E \setminus G)$ *and satisfies the following "conservation equations":*

c) $(r_t K)^\perp = \beta_t \geq 0$ *on* H*,*

d) $(r_t K)^\perp = -\int_\cdot j_t^{\mathrm{out}} \,\mathrm{d}\mathfrak{s} \geq 0$ *on* $\partial E \setminus G$*.*

See Appendix E for the proof.

6. Examples

6.1. A class of models with spontaneous jumps

Our first series of examples covers a large family of models without forced jumps ($G = \varnothing$). The reset kernel K is assumed to satisfy the following assumption:

Assumption 14. *There exists a kernel* K^* *on* E *such that*

$$\mathfrak{m}(\mathrm{d}x)\, K(x, \mathrm{d}y) = \mathfrak{m}(\mathrm{d}y)\, K^*(y, \mathrm{d}x)\,. \tag{24}$$

(We do *not* assume that K^* is a Markov kernel, i.e. that $K^*(y,\cdot)$ is a probability measure.) The following result is an easy consequence of Theorem 8:

Corollary 15. *If there exists a pdf $p \in C^{2,1}(E \times \mathbb{R}_+)$, then the measures r_t and r_t^{src} are absolutely continuous with respect to $\mathfrak{m}$,*

$$\frac{\mathrm{d}r_t}{\mathrm{d}\mathfrak{m}} \;=\; \lambda\, p_t\,, \quad \frac{\mathrm{d}r_t^{\mathrm{src}}}{\mathrm{d}\mathfrak{m}} \;=\; K^*\left(\lambda\, p_t\right)\,, \tag{25}$$

and the following evolution equation holds:

$$\frac{\partial p_t}{\partial t} \;=\; L^* p_t \,+\, K^*\left(\lambda\, p_t\right) \,-\, \lambda\, p_t\,. \tag{26}$$

Assumption 14 holds for several classes of models known in the literature: pure jump processes with an absolutely continuous reset kernel, the switching diffusions of Ghosh et al. [19, 18] and also the SHS of Hespanha [20].

Example 16. Pure jump processes occur when $L = 0$, i.e. when there is no continuous dynamics. We consider here the case where K is absolutely continuous: $K(x, \mathrm{d}y) = k(x, y)\, \mathfrak{m}(\mathrm{d}y)$. For instance, if the amplitude of the jumps is independent of the pre-jump state and distributed according to the pdf ρ, then $k(x, y) = \rho(y - x)$. In this case Assumption 14 holds with $K^*(x, \mathrm{d}y) = k(y, x)\, \mathfrak{m}(\mathrm{d}y)$. Introducing the function $\gamma(x, y) = \lambda(x)k(x, y)$, equation 26 turns into the well-known *master equation* [17, eq. 3.5.2]:

$$\frac{\partial p}{\partial t}(y, t) \;=\; \int \big(\gamma(x, y)p(x, t) - \gamma(y, x)p(y, t)\big)\, \mathfrak{m}(\mathrm{d}x)\,. \tag{27}$$

In particular, when all modes are purely discrete ($n_q = 0$), this is just the usual forward Kolmogorov equation for a continuous-time Markov chain.

Example 17. In the case of switching diffusions, the state space is of the form $E = \mathcal{Q} \times \mathbb{R}^n$ (with $\mathcal{Q}$ a countable set and $n \geq 1$) and the reset kernel of the form

$$K\big((q, z), \cdot\big) = \sum_{q' \neq q} \pi_{qq'}(z)\, \delta_{(q', z)}\,, \tag{28}$$

where $\pi(z) = (\pi_{qq'}(z))$ is a stochastic matrix for all $z \in \mathbb{R}^n$. Assumption 14 is fulfilled with K^* defined by

$$K^*\big((q, z), \cdot\big) = \sum_{q' \neq q} \pi_{q'q}(z)\, \delta_{(q', z)}\,. \tag{29}$$

Equation 15 becomes in this case the familiar generalized FPK equation for switching diffusion processes [see, e.g., 23, 24]: for all $x = (q, z) \in E$ and $t \geq 0$,

$$\frac{\partial p}{\partial t}(x, t) = (L^* p_t)(x) + \sum_{q' \neq q} \lambda_{q'q}(z)\, p_t(q', z) - \lambda(x)\, p_t(x)\,, \tag{30}$$

where $\lambda_{q'q}(z) = \lambda(q', z)\, \pi_{q'q}(z)$.

Example 18. The SHS of Hespanha [20] are also defined on $E = \mathcal{Q} \times \mathbb{R}^n$, but this time the post-jump state X_{τ_k} is determined by applying a reset map $\Psi : E \to E^0$ to the pre-jump state $X_{\tau_k}^-$, Ψ being chosen randomly in a finite of reset maps Ψ_k. The reset kernel can therefore be written as

$$K(x, \cdot) = \sum_k \pi_k(x)\, \delta_{\Psi_k(x)}\,, \tag{31}$$

with $\pi_k(x)$ the probability of choosing the reset map Ψ_k given that $X_{\tau_k}^- = x$. Provided that the functions Ψ_k are local C^1-diffeomorphisms, the kernel K fulfills Assumption 14 with

$$K^*(x, \cdot) = \sum_k \sum_{y \in \Psi_k^{-1}(\{x\})} \pi_k(y)\, \big|J_k(y)\big|^{-1}\, \delta_y\,, \tag{32}$$

where $J_k(y)$ is the Jacobian determinant of Ψ_k at y. Therefore, introducing a stochastic intensity $\lambda_k = \lambda\, \varrho_k$ for each one of the reset maps, we recover thanks to Corollary 15 the generalized FPK equation given by Hespanha [20, p. 1364]:

$$\frac{\partial p}{\partial t}(x, t) \;=\; (L^* p_t)(x) + \sum_k \sum_{y \in \Psi_k^{-1}(\{x\})} \left(\frac{\lambda_k\, p_t}{|J_k|}(y) \,-\, (\lambda_k\, p_t)(x) \right)\,. \tag{33}$$

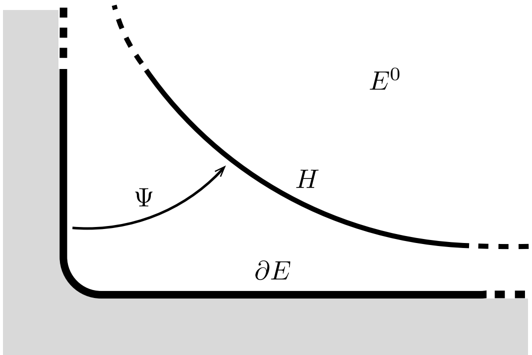

Figure 1: State space for the model of subsection 6.2.

6.2. A class of models with forced jumps

The measure-valued formulation of the generalized FPK equation (14) paves the way for an easier proof of some recent results [3], concerning GSHS with forced jumps and deterministic resets. A typical example of this class of process is the thermostat model of Malhamé and Chong [28], which has been extended to several dimensions in [3]. We consider the class of GSHS models satisfying the following assumptions.

Assumption 19. *a) The model only has forced jumps ($\lambda = 0$) with deterministic resets, i.e. there exists a map $\Psi : G \to E^0$ such that $K(x, \cdot\,) = \delta_{\Psi(x)}$ for all $x \in G$.*

b) All modes have the same dimension $n_q = n$, the guard set is the whole boundary ($G = \partial E$) and is of class C^1 (in particular, it has no corners).

c) $H = \Psi(G)$ is a C^1 hypersurface, closed in E, and Ψ is a C^1-diffeomorphism from G to H.

The state space of this model is depicted on figure 1. The assumption that $G = \partial E$ is only here for the sake of simplicity and could easily be relaxed. Boundaries with corners and piecewise smooth reset maps could be considered as well. The model considered in [3] also includes purely discrete modes (i.e., $n_q \in \{0, n\}$ for each $q \in \mathcal{Q}$), which cause no real additional difficulty.

The measure $K(x, \cdot\,)$ is supported by H for all x, which implies that the source term $r_t K$ is also supported by H, hence is $\mathfrak{m}$-singular. Therefore, even if the diffusion is non-degenerate (i.e. the diffusion matrix (a^{ij}) is uniformly positive definite), we know from subsection 5.3 that the pdf p_t will not be smooth on H. Accordingly, we make the following smoothness assumption for the measures μ_t:

Assumption 20. *a) μ_t admits a pdf p_t with respect to $\mathfrak{m}$ on the whole state space, for all $t \geq 0$,*

b) $p \in C^{2,1}\left((E \setminus H) \times \mathbb{R}_+\right)$, p_t and ∇p_t have at most a jump discontinuity (discontinuity of the first kind) on H.

Then Assumption 12 holds, which allows Theorem 13 to be applied. Moreover, the result of subsection C.1 holds on each component C of $E \setminus H$ — i.e., for all $\varphi \in C^2_c(E)$,

$$(L^* \mu_t)\,(\varphi_{|C}) \;=\; \int_C \varphi\, F p_t \,\mathrm{d}\mathfrak{m} + \int_{\partial C} \varphi\, \langle \mathbf{j}_t, \mathbf{n}_{\partial C} \rangle \;\mathrm{d}\mathfrak{s} + \frac{1}{2} \sum_{l=1}^{r} \int_{\partial C} \mathbf{f}_l \varphi\, p_t \;\langle \mathbf{f}_l, \mathbf{n}_{\partial C} \rangle \;\mathrm{d}\mathfrak{s}\,,$$

where $\mathbf{n}_{\partial C}$ is the outward-pointing unit vector on ∂C. Summing over the components and using Proposition 11.b yields

$$(L^* \mu_t)\,\varphi \;=\; \int_E F p_t\, \varphi\, \mathrm{d}\mathfrak{m} \;+\; \int_G j_t^{\mathrm{out}}\, \varphi\, \mathrm{d}\mathfrak{s} \;-\; \int_H j_t^{\mathrm{in}}\, \varphi\, \mathrm{d}\mathfrak{s} \;-\; \frac{1}{2} \sum_{l=1}^{r} \int_H \mathbf{f}_l \varphi\; \langle \mathbf{f}_l, \mathbf{n}_{ab} \rangle \; \left(p_t^b - p_t^a\right)\, \mathrm{d}\mathfrak{s}_H\,, \tag{34}$$

where $\mathbf{n}_{ab}$ is the unit normal vector on H oriented from side a to side b and $j_t^{\mathrm{in}} = \left\langle \mathbf{j}_t^{(b)} - \mathbf{j}_t^{(a)},\, \mathbf{n}_{ab} \right\rangle$. The superscripts a/b indicate the value of a discontinuous function on the corresponding side of H (but none of these quantities actually depend on the chosen labelling of the sides). The last term on the right-hand side vanishes, because $L^*\mu_t$ is a Radon measure by Theorem 8, whereas this term only involves the first-order derivatives of φ (through $\mathbf{f}_l\varphi$). Therefore, p_t is continuous on the set $H^0 = \{x \in H,\ \exists l \in \{1,\dots,r\},\ \langle \mathbf{f}_l, \mathbf{n}_{ab}\rangle \neq 0\}$ and the measures β_t of Theorem 13 are given by

$$\beta_t(\Gamma) \;=\; \int_{H\cap\Gamma} j_t^{\mathrm{in}}\, \mathrm{d}\mathfrak{s}_H\,. \tag{35}$$

The conclusions of Theorem 13 can then be summarized as follows:

a) The usual Fokker-Planck equation, $\partial p_t/\partial t = F p_t$, holds on each component of $E^0 \setminus H$.

b) The conservation probability current through the reset map is ensured by the relation

$$j_t^{\mathrm{out}} \;=\; \frac{\mathrm{d}(\mathfrak{s}_H \circ \Psi)}{\mathrm{d}\mathfrak{s}}\; j_t^{\mathrm{in}} \circ \Psi\,, \tag{36}$$

where $\mathfrak{s}_H \circ \Psi$ is the pushforward (image measure) of $\mathfrak{s}_H$ by Ψ^{-1}.

c) The absorbing boundary condition, $p_t = 0$, holds on G^0.

d) The density p_t is continuous on H^0.

7. Conclusions

A general measure-theoretic formulation of the Fokker-Planck-Kolmogorov equation has been presented, in the modern framework of GSHSs. This formulation is new, and should allow GSHS practitioners to get a better understanding of how a given system behaves in terms of evolution of the probability mass in the state space.

Technical tools have been provided, in order to derive the explicit form of the evolution equation when a probability density function exists and satisfy sufficient regularity conditions. In particular, it has been shown that the general FPK equation allows to recover all previously known instances of the FPK equation for stochastic hybrid systems.

Of course, an important issue is now to provide sufficient conditions for the existence of a “smooth enough” probability density function and the uniqueness of the solution to the generalized FPK equation. The literature already provides such conditions for processes defined by stochastic differential equations and (in some cases) switching diffusions; see, e.g., [5, 22, 25] and the references therein. Extending these results to other types of GSHSs, including models with forced jumps, is an important perspective for future work.

In addition to providing a better understanding of GSHSs, the FPK equation is also a powerful tool for the analysis of low-dimensional systems, for which an approximate solution can be obtained using numerical methods (for instance finite volume methods). It is especially useful for the computation of the stationary distribution, as shown in [4] a nontrivial three-dimensional model of a wind turbine. This type of application of the FPK equation to the analysis of GSHSs relies on the availability of software components allowing an easy implementation of efficient numerical methods, in the spirit of Mitchell’s Level Set Toolbox [29, 30] for Hamilton-Jacobi equations. The development of such a toolbox is another important direction for future work.

References

[1] J. Bect. *Processus de Markov diffusif par morceaux: outils analytiques et numériques*. PhD thesis, Université Paris Sud XI, 2007. In french.

[2] J. Bect. A unifying formulation of the Fokker-Planck-Kolmogorov equation for general stochastic hybrid systems. In *Proceedings of the 17th IFAC World Congress*, 2008.

[3] J. Bect, H. Baili, and G. Fleury. Generalized Fokker-Planck equation for piecewise-diffusion processes with boundary hitting resets. In *Proceedings of the 17th International Symposium on the Mathematical Theory of Networks and Systems (MTNS 2006)*, Kyoto, Japan, 2006.

[4] J. Bect, Y. Phulpin, H. Baili, and G. Fleury. On the Fokker-Planck equation for stochastic hybrid systems: application to a wind turbine model. In *Proc. of the 9th Int. Conf. on Probabilistic Methods Applied to Power Systems (PMAPS)*, 2006.

[5] K. Bichteler, J.-B. Gravereaux, and J. Jacod. *Malliavin Calculus for Processes with Jumps*. Gordon and Breach Science Publishers, 1987.

[6] R. M. Blumenthal and R. K. Getoor. *Markov Processes and Potential Theory*. Academic Press, New York, 1968. Pure and Applied Mathematics, Vol. 29.

[7] M. L. Bujorianu and J. Lygeros. General stochastic hybrid systems: modelling and optimal control. In *Proceedings of the 43rd IEEE Conference on Decision and Control (CDC 2004)*, volume 2, pages 1872–1877, 2004.

[8] M. L. Bujorianu and J. Lygeros. Toward a general theory of stochastic hybrid systems. In *Stochastic Hybrid Systems: Theory and Safety Critical Applications*, volume 337 of *LNCIS*, pages 3–30. Springer Verlag, 2006.

[9] M. L. Bujorianu, J. Lygeros, W. Glover, and G. Pola. A stochastic hybrid system modeling framework. Technical Report WP1, Deliv. 1.2, HYBRIDGE (IST-2001-32460), 2003.
[10] D. L. Cohn. *Measure Theory*. Birkhäuser, 1980.
[11] M. H. A. Davis. Piecewise deterministic Markov processes: a general class of nondiffusion stochastic models. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 46:353–388, 1984.
[12] M. H. A. Davis. *Markov Models and Optimization*. Chapman & Hall, London, 1993.
[13] E. B. Dynkin. *Markov Processes*. Springer, 1965.
[14] M. H. C. Everdij and H. A. P. Blom. Hybrid Petri nets with diffusion that have into-mappings with generalised stochastic hybrid processes. *Lecture Notes in Control and Information Sciences*, 337:31–63, 2006.
[15] W. Feller. The parabolic differential equations and the associated semi-groups of transformations. *Annals of Mathematics*, 55(3):468–519, 1952.
[16] W. Feller. Diffusion processes in one dimension. *Transactions of the American Mathematical Society*, 77:1–31, 1954.
[17] C. W. Gardiner. *Handbook of Stochastic Methods for Physics, Chemistry and the Natural Sciences*. Springer-Verlag, 1985.
[18] M. K. Ghosh, A. Arapostathis, and S. I. Marcus. Optimal control of switching diffusions with application to flexible manufacturing systems. *SIAM Journal on Control and Optimization*, 30(6):1–23, 1992.
[19] M. K. Ghosh, A. Arapostathis, and S. I. Marcus. Ergodic control of switching diffusions. *SIAM Journal on Control and Optimization*, 35(6): 1952–1988, 1997.
[20] J. P. Hespanha. A model for stochastic hybrid systems with application to communication networks. *Nonlinear Analysis: Theory, Methods and Applications*, 62:1353–1383, 2005.
[21] J. Hu, J. Lygeros, and S. Sastry. Towards a theory of stochastic hybrid systems. pages 160–173. Springer, 2000.
[22] N. Ikeda and S. Watanabe. *Stochastic differential equations and diffusion processes*. North Holland-Kodansha, Tokyo, 1981.
[23] V. Kontorovich and V. Lyandres. Dynamic systems with random structure: an approach to the generation of nonstationary stochastic processes. *Journal of the Franklin Institute*, 336:939–954, 1999.
[24] J. Krystul, A. Bagchi, and H. A. P. Blom. Risk decomposition and assessment methods. Technical Report D8.1, HYBRIDGE (IST-2001-32460), 2003.
[25] C. Le Bris and PL Lions. Existence and uniqueness of solutions to Fokker–Planck type equations with irregular coefficients. *Communications in Partial Differential Equations*, 33(7):1272–1317, 2008.
[26] J. M. Lee. *Introduction to smooth manifolds*. Number 218 in Graduate Texts in Mathematics. Springer-Verlag, 2003.
[27] I. Lubashevsky, R. Friedrich, R. Mahnke, A. Ushakov, and N. Kubrakov. Boundary singularities and boundary conditions for the Fokker-Planck equations, 2006.
[28] R. Malhamé and C. Chong. Electric load model synthesis by diffusion approximation of a high-order hybrid-state stochastic system. *IEEE Transactions on Automatic Control*, 30(9):854–860, 1985.
[29] I. M. Mitchell. The flexible, extensible and efficient toolbox of level set methods. *Journal of Scientific Computing*, 35(2):300–329, 2008.
[30] I. M. Mitchell and J. A. Templeton. A toolbox of Hamilton-Jacobi solvers for analysis of nondeterministic continuous and hybrid systems. volume 3414, pages 480–494. Springer, 2005.
[31] G. Pola, M. L. Bujorianu, J. Lygeros, and M. D. Di Benedetto. Stochastic hybrid models: an overview with application to air traffic management. In *ADHS 03, IFAC Conference on Analysis and Design of Hybrid Systems*, June 2003.
[32] L. C. G. Rogers and D. Williams. *Diffusions, Markov Processes, and Martingales. Volume 2: Itô Calculus. 2nd edition.* Cambridge University Press, 2000.
[33] W. Rudin. *Function Analysis*. McGraw-Hill, 1973.
[34] M. Sharpe. *General Theory of Markov Processes*. Academic Press, 1988.
[35] J. B. Walsh and M. Weil. Représentation de temps terminaux et applications aux fonctionnelles additives et aux systèmes de Lévy. *Annales scientifiques de l'E.N.S. 4ᵉsérie*, 5(1):121–155, 1972.

A. Smooth maps and vector fields on E

The following definitions are natural extensions to the hybrid state space E of the usual definition on subsets of $\mathbb{R}^n$.

A map $\varphi : E \to \mathbb{R}$ is said to be k-times continuously differentiable on E — in short, $\varphi \in C^k(E)$ — if $\varphi_q = \varphi(q,\,\cdot\,)$ is C^k on E_q in the usual sense for all $q \in \mathcal{Q} \setminus \mathcal{Q}^{\mathrm{d}}$, i.e. if there is an open subset U of $\mathbb{R}^{n_q}$ such that $E_q \subset U$ and φ extends to a C^k map on U.

A vector field $\mathbf{g}$ on E is defined as a first-order differential operator with respect to the continuous variables. Its action on a continuously differentiable function $\varphi \in C^1(E)$ will be denoted by $\mathbf{g}\varphi$, where

$$(\mathbf{g}\varphi)(q,z) = \begin{cases} \sum_{i=1}^{n_q} \mathbf{g}^i(q,z) \frac{\mathrm{d}\varphi}{\mathrm{d}z^i}(q,z) & \text{on } E \setminus E^{\mathrm{d}}\,, \\ 0 & \text{on } E^{\mathrm{d}}\,. \end{cases} \tag{37}$$

The number of “components” of $\mathbf{g}$ depends on the mode q. To simplify the notations, we shall agree that the indices i and j always correspond to summations on the number of continuous variables, and drop the explicit dependence on q. For instance, the definition of $\mathbf{g}\varphi$ on $E \setminus E^{\mathrm{d}}$ can be rewritten as $\mathbf{g}\varphi = \sum_i \mathbf{g}^i \frac{\partial\varphi}{\partial z^i}$. A vector field is said to be k-times continuously differentiable on E if $\mathbf{g}^i(q,\,\cdot\,)$ is C^k on E_q in the usual sense for all $q \in \mathcal{Q} \setminus \mathcal{Q}^{\mathrm{d}}$ and all $i \in \{1,\dots,n_q\}$.

Finally, under Assumption 10, the following version of the divergence theorem holds (since each component of the state space is a C^2-manifold with corners) :

Theorem 21 (see, e.g., [26])**.** *Let Assumption 10 hold. Then, for all compactly supported C^1 vector field $\mathbf{f}$ on E,*

$$\int_E \mathrm{div}(\mathbf{f})\,\mathrm{d}\mathfrak{m} \;=\; \int_{\partial E} \langle \mathbf{f}, \mathbf{n} \rangle \;\mathrm{d}\mathfrak{s}\,. \tag{38}$$

B. Proof of Theorem 8

Let $C^2_{\mathrm{c}}(E)$ denote the set of all compactly supported $\varphi \in C^2(E)$. The following lemma is an easy consequence of the smoothness of the vector fields:

Lemma 22. *For all $\varphi \in C^2(E)$, $t \mapsto \int_0^t (L^*\mu_s)(\varphi)\,\mathrm{d}s$ is differentiable on the right, with the right-continuous derivative $t \mapsto (L^*\mu_t)(\varphi)$.*

In the sequel, "right-continuous" is abbreviated as "rc".

$\diamond$ Assume that both 8.a and 8.b hold. Then each term of (13) has a t-derivative on the right. Differentiating both sides proves that (14) holds for all $t \geq 0$, hence that $L^*\mu_t$ is a Radon measure and that $t \mapsto L^*\mu_t$ is rc. Moreover, integrating the inequality $|L^*\mu_t| \leq |\mu'_t| + 2r_t$ yields that, for all $\Gamma \in \mathcal{E}_{\mathrm{c}}$,

$$\int_0^t |L^*\mu_s|\,(\Gamma)\,\mathrm{d}s \leq \int_0^t |\mu'_s|\,(\Gamma)\,\mathrm{d}s + 2\,\boldsymbol{E}\big\{N_t\big\} \leq +\infty\,. \tag{39}$$

Therefore $t \mapsto |L^*\mu_s|$ is locally integrable, which proves 8.c.

$\diamond$ Assume now that 8.a and 8.c hold, and set $\mu'_t = L^*\mu_t + r_t(K - I)$, for all $t \geq 0$. Clearly, μ'_t is a Radon measure, $t \mapsto \mu'_t$ is rc and

$$\int_0^t \mu'_t\varphi = (\mu_t - \mu_0)\varphi\,, \quad \forall t \geq 0\,, \quad \forall \varphi \in C^2_{\mathrm{c}}(E)\,. \tag{40}$$

Moreover, for all $\Gamma \in \mathcal{E}_{\mathrm{c}}$,

$$\int_0^t |\mu'_s|\,(\Gamma)\,\mathrm{d}s \leq \int_0^t |L^*\mu_s|\,(\Gamma)\,\mathrm{d}s + 2\,\boldsymbol{E}\big\{N_t\big\} \leq +\infty\,, \tag{41}$$

which shows that $t \mapsto |\mu'_s|$ is locally integrable. Therefore, using standard approximation techniques and a monotone class argument, it can be proved that (40) still holds for $\varphi = \mathbb{1}_\Gamma$, $\Gamma \in \mathcal{E}_{\mathrm{c}}$, i.e. that $t \mapsto \mu'_t$ is the "derivative" of $t \mapsto \mu_t$ in the sense of definition 7.b.

$\diamond$ Finally, assume that 8.b and 8.c hold. Then, for all $\varphi \in C^2_{\mathrm{c}}(E)$, equation (13) can be rewritten as

$$\iint\limits_{G\times]0;t]} \varphi(x)\;\big(R^G(\mathrm{d}x,\mathrm{d}s) - (L^*\mu_s)(\mathrm{d}x)\mathrm{d}s\big) \;=\; \iint\limits_{E^0\times]0;t]} \varphi(x)\;\big((R^G K)(\mathrm{d}x,\mathrm{d}s) - \xi_s(\mathrm{d}x)\,\mathrm{d}s\big)\,, \tag{42}$$

where $\xi_s = \mu'_s - \big(L^*\mu_s\big)(E^0 \cap \cdot\,) - r^0(K - I)$. The measures R^G and r^0 have been defined in subsection 3.1. Clearly, $\xi_t \in \mathcal{M}_{\mathrm{c}}(E)$ and $t \mapsto \xi_t$ is locally integrable. Using once more standard approximation techniques, one can prove that (42) still holds when $\varphi = \mathbb{1}_\Gamma$, with Γ a compact subset of G. In this case the right-hand side vanishes, yielding

$$R^G(\Gamma\times]0;t]) \;=\; \int_0^t (L^*\mu_s)(\Gamma)\,\mathrm{d}s\,. \tag{43}$$

Moreover, since $t \mapsto R^G(\Gamma\times]0;t])$ is increasing and $t \mapsto (L^*\mu_t)(\Gamma)$ is rc, we have $(L^*\mu_t)(\Gamma) \geq 0$ for all $t \geq 0$. This allows to extend (42) to all $\Gamma \in \mathcal{E}_{\mathrm{c}}$, using a monotone class argument, thus proving the existence of a mean jump intensity $r^G_t = (L^*\mu_s)(G \cap \cdot\,)$ for the forced jumps.

C. Proof of Proposition 11

Since p is of class $C^{2,1}$ on $U \times \mathbb{R}_+$, it is easily seen that the assertions 8.b and 8.c hold on U, with $\mu'_t(\mathrm{d}x) = \frac{\partial p}{\partial t}(x,t)\,\mathfrak{m}(\mathrm{d}x)$. Using the same arguments as in the proof of Theorem 8, it follows that 8.a and the generalized FPK equation hold on U as well. The rest of the proof is split into three parts.

C.1. Computation of $L^*\mu_t$

Using the definitions of L (equation (4)) and $\mathbf{j}_t$ (equation (16)), we find that

$$\begin{aligned}
L\varphi\, p_t \;&=\; \left(\sum_i \mathbf{f}_0^i\,\frac{\partial\varphi}{\partial z^i} + \frac{1}{2}\sum_{i,j} a^{ij}\,\frac{\partial^2\varphi}{\partial z^i\partial z^j}\right)\; p_t \\
&=\; \mathbf{j}_t\varphi + \frac{1}{2}\sum_{i,j}\frac{\partial(a^{ij}p_t)}{\partial z^j}\,\frac{\partial\varphi}{\partial z^i} + \sum_{i,j} a^{ij}p_t\frac{\partial^2\varphi}{\partial z^i\partial z^j} \\
&=\; \mathbf{j}_t\varphi + \frac{1}{2}\sum_{i,j}\frac{\partial}{\partial z^j}\left(a^{ij}p_t\frac{\partial\varphi}{\partial z^i}\right) \\
&=\; \mathbf{j}_t\varphi + \frac{1}{2}\sum_{l=1}^{r}\sum_{i,j}\frac{\partial}{\partial z^j}\left(\mathbf{f}_l^i\mathbf{f}_l^j p_t\frac{\partial\varphi}{\partial z^i}\right) \\
&=\; \mathbf{j}_t\varphi + \frac{1}{2}\sum_{l=1}^{r}\operatorname{div}\left(\mathbf{f}_l\varphi\, p_t\,\mathbf{f}_l\right)\;. \qquad (44)
\end{aligned}$$

Moreover, using the product rule for the divergence operator and the fact that $Fp_t = -\operatorname{div}(\mathbf{j}_t)$, which is a direct consequence of equations (16) and (21), we get

$$\mathbf{j}_t\varphi \;=\; \operatorname{div}(\varphi\mathbf{j}_t) \,-\, \varphi\,\operatorname{div}(\mathbf{j}_t) \;=\; \operatorname{div}(\varphi\mathbf{j}_t) \,+\, \varphi\,Fp_t\,. \qquad (45)$$

Finally, equations (44) and (45) together with the divergence theorem yield :

$$\begin{aligned}
(L^*\mu_t)\,(\varphi) \;&=\; \int_E L\varphi\,\mathrm{d}\mu_t \;=\; \int_E L\varphi\,p_t\,\mathrm{d}\mathfrak{m} \\
&=\; \int_E\left(\operatorname{div}(\varphi\mathbf{j}_t) \,+\, \varphi\,Fp_t + \frac{1}{2}\sum_{l=1}^{r}\operatorname{div}\left(\mathbf{f}_l\varphi\,p_t\,\mathbf{f}_l\right)\right)\,\mathrm{d}\mathfrak{m} \\
&=\; \int_{\partial E}\langle\varphi\mathbf{j}_t,\mathbf{n}\rangle\;\mathrm{d}\mathfrak{s} + \int_E \varphi\,Fp_t\,\mathrm{d}\mathfrak{m} + \frac{1}{2}\sum_{l=1}^{r}\int_{\partial E}\langle\mathbf{f}_l\varphi\,p_t\,\mathbf{f}_l,\mathbf{n}\rangle\;\mathrm{d}\mathfrak{s} \\
&=\; \int_E \varphi\,Fp_t\,\mathrm{d}\mathfrak{m} + \int_{\partial E}\varphi j_t^{\mathrm{out}}\,\mathrm{d}\mathfrak{s} + \frac{1}{2}\sum_{l=1}^{r}\int_{\partial E}\mathbf{f}_l\varphi\,p_t\;\langle\mathbf{f}_l,\mathbf{n}\rangle\;\mathrm{d}\mathfrak{s}\,. \qquad (46)
\end{aligned}$$

C.2. Proof of assertion 11.b

Let $V = (\partial E)_{\text{smooth}} \cap U$, where $(\partial E)_{\text{smooth}}$ denote the smooth part of the boundary. V is an open subset of ∂E. For each $\eta \in C^2_{\mathrm{c}}(V)$, there exists a sequence of functions $\varphi_n \in C^2_{\mathrm{c}}(U)$, with their support in a fixed compact set, such that $\varphi_n = 0$ and $\partial\varphi_n/\partial\mathbf{n} = \eta$ on V, and $\varphi_n \to 0$ uniformly. Equation (46) holds for each n. Taking limits with respect to n on both sides, and using the fact that $(L^*\mu_t)\,(\varphi_n) \to 0$ (since $L^*\mu_t$ is a Radon measure on U), we find that $\sum_l \int_{\partial E}\eta\,p_t\;\langle\mathbf{f}_l,\mathbf{n}\rangle^2\;\mathrm{d}\mathfrak{s} = 0$. Therefore, $x \mapsto p(x,t)\;\sum_l\langle\mathbf{f}_l,\mathbf{n}\rangle^2$ vanishes $\mathfrak{s}$-almost everywhere on V, hence everywhere on ∂E by continuity. This proves assertion 11.b since $\sum_l\langle\mathbf{f}_l,\mathbf{n}\rangle^2 > 0$ on G^0.

C.3. Proof of assertion 11.a

It is now proved that, for each $x \in \partial E$, either $p_t(x) = 0$ or $\langle\mathbf{f}_l,\mathbf{n}\rangle_x = 0$ for all $l \in \{1,\ldots,r\}$. As a consequence, for all $\varphi \in C^2_{\mathrm{c}}(U)$, the last term of equation (46) vanishes :

$$(L^*\mu_t)\,(\varphi) \;=\; \int_E Fp_t\,\varphi\,\mathrm{d}\mathfrak{m} \,+\, \int_{\partial E}\varphi j_t^{\mathrm{out}}\,\mathrm{d}\mathfrak{s}\,, \qquad (47)$$

and therefore the Radon measure $L^*\mu_t$ can be rewritten as

$$(L^*\mu_t)\,(\Gamma) \;=\; \int_\Gamma Fp_t\,\mathrm{d}\mathfrak{m} \,+\, \int_{\partial E\cap\Gamma} j_t^{\mathrm{out}}\,\mathrm{d}\mathfrak{s} \qquad (48)$$

for all $\Gamma \in \mathcal{E}_\mathrm{c}$ such that $\Gamma \subset U$. Substituting the result into the generalized FPK equation (14) yields

$$\mu_t'(\Gamma) \;=\; \int_\Gamma F p_t \,\mathrm{d}\mathfrak{m} \;+\; \int_{\partial E \cap \Gamma} j_t^{\mathrm{out}} \,\mathrm{d}\mathfrak{s} + (r_t K)(\Gamma) - r_t(\Gamma)\,. \tag{49}$$

Finally, for all $t \geq 0$ and all $\Gamma \in \mathcal{E}_\mathrm{c}$ such that $\Gamma \subset G \cap U$, we have $\mu_t'(\Gamma) = 0$ and $\int_\Gamma F p_t \,\mathrm{d}\mathfrak{m} = 0$, both because $\mathfrak{m}(\Gamma) = 0$, and $(r_t K)(\Gamma) = 0$ because K is a kernel from E to E^0. Therefore, as a consequence of (49), $r_t^G(\Gamma) = \int_\Gamma j_t^{\mathrm{out}} \,\mathrm{d}\mathfrak{s}$ for all $\Gamma \in \mathcal{E}_\mathrm{c}$ such that $\Gamma \subset G \cap U$. The outward current j_t^{out} is thus positive $\mathfrak{s}$-almost everywhere on U, hence everywhere by continuity. This proves assertion 11.a.

D. The Lebesgue-Radon-Nikodym theorem

This appendix recalls a fundamental result of measure theory, which is used in the statement and proof of Theorem 13. The reader is referred to, e.g., [10, chapter 4], for basic definitions and terminology not recalled here. In this section, $(E, \mathcal{E})$ denotes any measurable space — not necessarily the hybrid state space as defined in subsection 2.1.

Theorem 23 (Lebesgue-Radon-Nikodym)**.** *Let ν_1, ν_2 be positive σ-finite measures on $(E, \mathcal{E})$. Then there is a unique positive σ-finite measure $\nu_1^\perp$ and a unique (up to ν_2-everywhere equality) measurable function f such that*

a) $\nu_1^\perp$ and ν_2 are mutually singular, i.e. $\exists A \in \mathcal{E}$, $\nu_1^\perp(A) = 0$ and $\nu_2(E \setminus A) = 0$,

b) ν_1 has the following decomposition

$$\nu_1(A) = \int_A f \,\mathrm{d}\nu_2 + \nu_1^\perp(A), \qquad \forall A \in \mathcal{E}\,. \tag{50}$$

The measures $A \mapsto \int_A f \,\mathrm{d}\nu_2$ and $\nu_1^\perp$ are respectively called the *absolutely continuous part* and the *singular part* of ν_1 with respect to ν_2. Equation (50) is called the Lebesgue-Radon-Nikodym decomposition of ν_1 with respect to ν_2. The function f is called the Radon-Nikodym derivative of ν_1 with respect to ν_2, is usually denoted by $\frac{\mathrm{d}\nu_1}{\mathrm{d}\nu_2}$.

E. Proof of Theorem 13

The assumption 8.b of Theorem 8 holds with $\mu'(\Gamma) = \int_\Gamma \frac{\partial p}{\partial t}$, since $\frac{\partial p}{\partial t}$ exists $\mathfrak{m}$-almost everywhere and is locally bounded. Moreover, as in the proof of Proposition 11, the existence of r_t^G on $G \cap U = G$ follows from the fact that 8.b and 8.c hold on U. We have thus proved that the assumptions 8.a (existence of r_t) and 8.b of Theorem 8 (existence of μ') hold on the whole state space, which implies that 8.c and the generalized FPK equation (14) hold as well.

As a consequence of equation (48), we have

$$\left(L^* \mu_t\right)(\Gamma \cap U) \;=\; \int_{\Gamma \cap U} F p_t \,\mathrm{d}\mathfrak{m} \;+\; \int_{\partial E \cap \Gamma \cap U} j_t^{\mathrm{out}} \,\mathrm{d}\mathfrak{s} \;=\; \int_\Gamma F p_t \,\mathrm{d}\mathfrak{m} \;+\; \int_{\partial E \cap \Gamma} j_t^{\mathrm{out}} \,\mathrm{d}\mathfrak{s} \tag{51}$$

for all $\Gamma \in \mathcal{E}_\mathrm{c}$, since $\mathfrak{m}(E \subset U) = 0$ and $\partial E \subset U$. Equation (22) thus simplifies into $\beta_t(\Gamma) = -\left(L^* \mu_t\right)(\Gamma \cap H)$, which proves that the measures β_t are supported by H.

According to (51), the generalized FPK equation (14) can be decomposed as

$$\mu_t'(\Gamma) \;=\; \int_\Gamma F p_t \,\mathrm{d}\mathfrak{m} \;+\; \int_{\partial E \cap \Gamma} j_t^{\mathrm{out}} \,\mathrm{d}\mathfrak{s} \;+\; \left(L^* \mu_t\right)(\Gamma \cap H) \;+\; r_t(K - I)(\Gamma)\,. \tag{52}$$

Uniqueness of the Lebesgue-Radon-Nikodym decomposition with respect to $\mathfrak{m}$ yields

$$\frac{\partial p}{\partial t} \;=\; F\, p_t \;+\; \frac{\mathrm{d}(r_t K)}{\mathrm{d}\mathfrak{m}} \;-\; \lambda\, p_t\,, \tag{53}$$

$$0 \;=\; \int_{\partial E \cap \Gamma} j_t^{\mathrm{out}} \,\mathrm{d}\mathfrak{s} \;+\; \left(L^* \mu_t\right)(\Gamma \cap H) \;+\; \left(r_t K\right)^\perp(\Gamma) \;-\; \left(r_t^G\right)(\Gamma)\,, \tag{54}$$

where we have used that

$$\begin{aligned}
\mu_t'(\mathrm{d}x) &= \frac{\partial p}{\partial t}(x,t)\,\mathfrak{m}(\mathrm{d}x)\,, \\
(r_tK)(\mathrm{d}x) &= \frac{\mathrm{d}(r_tK)}{\mathrm{d}\mathfrak{m}}(x)\,\mathfrak{m}(\mathrm{d}x) + (r_tK)^{\perp}(\mathrm{d}x)\,, \\
r_t(\mathrm{d}x) &= \lambda(x)\,p_t(x)\,\mathfrak{m}(\mathrm{d}x) + r_t^G(\mathrm{d}x)\,.
\end{aligned}$$

The first line of the system (53)–(54) is precisely (23), and the second one readily splits into 13.c and 13.d by considering the terms that are supported respectively by H and $\partial E \setminus G$. The proof of Theorem 13 is thus complete.